

Detection of polluting plumes ejected from NYC buildings

Ben Steers¹, Jon Kastelan¹, Chun Chieh Tsai¹, Federica B Bianco², Gregory Dobler², and CUSP capstone manager³

¹New York University (NYU)

²NYU Center for Urban Science & Progress

³Affiliation not available

April 28, 2020



2018 Capstones

As part of the urban metabolism, city buildings consume resources and use energy, producing environmental impacts on the surrounding air by emitting plumes of pollution. Plumes that have been observed in Manhattan range from water vapor emitted from heating and cooling systems' steam vents to CO₂ and dangerous chemical compounds (e.g. ammonia, methane). City agencies are interested in detecting and tracking these plumes as they provide evidence for signs of urban activity, cultivation of living and working spaces and can support the provision of services whilst monitoring environmental impacts. The Urban Observatory at New York University's Center for Urban Science and Progress ([CUSP-UO](#)) continuously images the Manhattan skyline at 0.1 Hz, and day-time images can be used to detect and characterize plumes from buildings in the scene. This project built and trained a deep convolutional neural network for detection and tracking of these plumes in near real-time. The project created a large training set of over 1,100 actual plumes as well as sources of contamination such as clouds, shadows and lights, and applied the

relevant network architecture for training of the model. The trained convolutional neural network was applied to the archival Urban Observatory data between two time periods: 26th October-31st December 2013 and 1st January-13th March 2015 to generate detections of building plume activity during those time periods. Buildings with high plume ejection rates were identified, and all plumes could be classified by their color (i.e. carbon vs water vapor). The final result was a detection of plumes emitted during the time periods that the dataset spans.

1 Introduction

Smoke plumes contribute significantly to the air pollution in cities posing serious health risks including heart disease, lung cancer, and asthma. The New York City Department of Health recently estimated that up to 2,700 premature deaths a year could be attributed to fine particulate matter and ozone in the air ([NYC Health, 2013](#); [Mills et al., 2008](#)). In 2005, it was estimated that approximately 10,000 buildings in the city burned number 4 and 6 heating oils, which emit more air polluting Particulate Matter (PM) 2.5, Sulphur Dioxide (SO₂) and nickel than alternative fuels [NYC Health \(2013\)](#). In 2007, NYC launched a sustainability program, titled PlaNYC, which aims to bring significant emission reductions, with a goal of 30%, by 2030 ([City of New York, 2013](#)). According to [City of New York \(2010\)](#), NYC buildings account for 75% of all of greenhouse gas emissions (including CO₂), meaning that in order to enact the necessary change, building energy usage needs to be addressed and understood. The wider implications of this study could impact many different city agencies and departments such as those overseeing en-

ergy, environment, health, transport, buildings, and housing.

The traditional methods for detection of plumes rely on extensive groups of connected sensors (sensor networks) that provide a local measure of air quality by detecting the presence of particles in the air. (Brink and Pebesma, 2013) This would require a dense and extensive network in order to gain a comprehensive view of the entire city, requiring permission from building owners to use their buildings, and a team to maintain the network. Additionally, all nodes need network connectivity to support data transfer, making the process cumbersome and costly. This project’s objective is to make the spatio-temporal tagging of the plumes a real-time and viable process using image data.

Automatic image-based smoke detection models from the literature span a variety of different methods, many using hand-engineered features (e.g. threshold setting). Çelik et al. (2007) take a statistical approach, using color models to detect both regions with smoke and those with fire which are constructed using hand-engineered color features such as setting thresholds for the color range. Yuan (2008) attempts to improve the false alarm rate of video-based smoke detection algorithms by incorporating the orientation of the smoke’s motion, helping remove the disturbance of other moving objects. Gubbi et al. (2009); Ko et al. (2013) use visual codebook style representations to detect the presence of smoke, employing support vector machines (SVMs) and random forest classifiers, respectively.

Recognizing that the existing literature was primarily rule-based models and hand-engineered features, and the potential of Convolutional Neural Networks (CNNs) given their demonstrated success in

image classification, [Karpathy et al. 2014](#). [Frizzi et al. \(2016\)](#) trained CNNs to detect fire and smoke in still images. Convolutional neural networks (CNNs) have emerged as the state-of-the-art image classification algorithm due to its efficient architecture that takes advantage of the stationarity and locality of patterns found in images and videos. Unlike other machine learning methods, Convolutional Neural Networks do not rely on engineered features, but rather extract the features most relevant for classification automatically based on a labeled training set. CNNs perform well for the goal of image classification; however what about the more detailed question of where a specific object is within an image. This may be the case for a plume (or multiple plumes) within an image. To correctly identify where the main objects in the images, Faster R-CNN is a specific form of CNN that includes a region proposal network which hypothesizes object locations via bounding boxes, and in a more efficient way than its predecessors, R-CNN and Fast R-CNN ([Ren et al., 2017](#)).

This project aims to create a method for detecting and recording plumes of pollution in NYC using images gathered from the Urban Observatory at New York University’s Center for Urban Science and Progress ([CUSP-UO](#)). The CUSP-UO studies the complex interactions between the physical, natural, and human components of the city as a coherent, definable system with the goal of enhancing public well-being, city operations, and future urban plans. CUSP-UO continuously images the Manhattan skyline at 0.1 Hz, for use in image based detection which is synoptic, persistent and non-intrusive. The daytime images can be used to detect and characterize plumes from buildings in the scene ([Dobler et al., 2015](#)). The project also aims to identify various statistics such as the origin, count and frequency

of the plumes. This will be performed by constructing a training dataset which will be used to train Faster R-CNN models for plume location. The locations can then be mapped to geographic coordinates to identify a given source building.

2 Data description

This project uses images from the CUSP-UO cityscape, collected every day across two distinct time periods: October 26th through December 31st 2013 and January 1st through March 13th 2015. 2160x4096 pixel RGB images with 8-bit depth are sampled every 10 seconds. The CUSP-UO has several camera deployments. The camera used in this paper is located in 1 Metrotech Center at an elevation of approximately 800 ft from the ground, pointing north. These images capture a diverse set of buildings from the Lower East Side up to Midtown of Manhattan (Figure 1).

The images will be divided into a training, testing and validation sets. Due to visibility concerns, only daytime images are considered, where daytime is defined to be between 5:00AM and 06:00PM year round. Most days comprise $\sim 8,630$ images in total, with $\sim 4,700$ daytime images. Some days have significantly less (Figure 2) and are likely periods where the camera had to be taken offline for maintenance. One day 2/9/2015 doesn't have any images. The camera was likely undergoing maintenance on that day. The total number of images the CUSP-UO is 1,133,811 with 612,902 during daytime hours.



Figure 1: Image from the CUSP-UO, a facility that continuously images the New York City skyline. This camera is located at 1 Metrotech Center facing north.

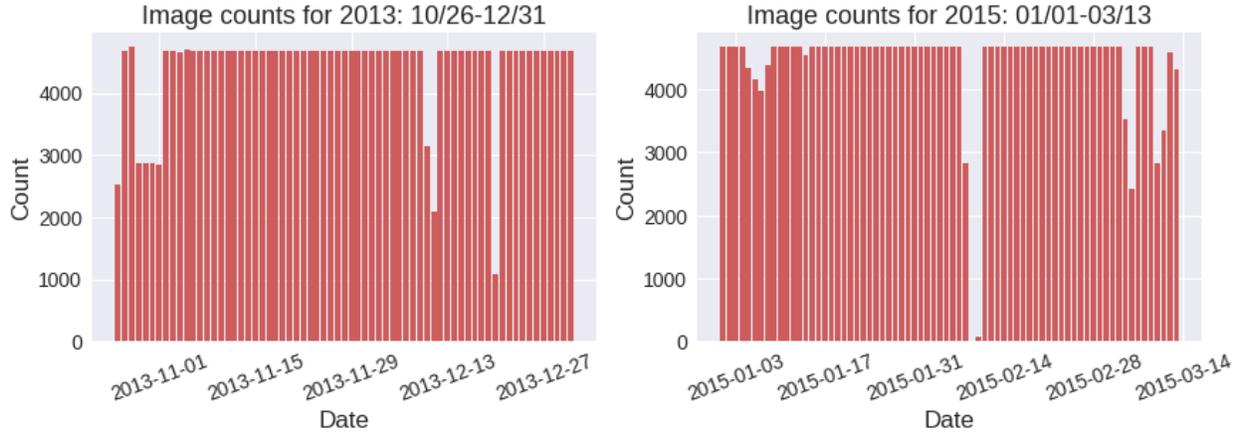


Figure 2: Number of images per day from the CUSP UO 1MTC location for the days being considered between 10/26-12/31 in 2013 and 01/01-03/13 in 2015. The typical day has 4,675 images per day between 5:00AM and 6:00PM.

3 Methodology

The approach taken for detection of polluting plumes in New York City was as follows:

1. **Background subtraction** - a differencing technique is applied to the images to remove stationary features which are common between images, and highlight differences between static and moving objects, including people, vehicles, clouds, shadows, and of course pollution plumes .
2. **Image labelling for compilation of training set** - using visual inspection of differenced images paired with statistical heuristics to identify images containing plumes along with other sources of noise and perturbation which a model may pick up as signal.
3. **Model development** - training and testing of Faster-RCNN model to develop a plume detection algorithm for use across the larger set of images (those not previously included in the training or testing set).
4. **Plume detection census** - run the model across the complete set of images to compile a census of all plumes along with statistics such as source, frequency, and type of plume emitted by buildings in the images.

4 Results

The Faster R-CNN model was trained on a down-sampled set of all tagged annotations (Table 1).

Class	Total	Train	Test	Validation
Plume	11618	5703	3480	2435
Light	400	186	133	81
Cloud	258	122	75	61
Shadow	158	80	44	34
Ambiguous	15	8	3	4

Table 1: Summary of down-sampled tagged annotations

The average precision (mAP) on the test set for detection of Plumes is 61.6%. Compared to the other classes, the model is best at detecting plumes with the highest mean AP, mean Recall and mean Precision. The mean Precision across all classes is relatively low, with mean Precision for plumes around 14.0%, whilst all other classes the mean Precision is less than 1%. The low mAP rates for the other classes is a result of the incomplete tagging. This was more severe with the sources of contamination because fewer examples were captured for each of those classes, both proportionally to the number of contaminants in the dataset and relative to the total plume count.

Class	Mean Average Precision	Mean Recall	Mean Precision
Plume	0.62	0.7	0.14
Light	0.06	0.23	0.0
Cloud	0.07	0.18	0.0
Shadow	0.24	0.52	0.0
Ambiguous	0.22	0.33	0.0

Table 2: Summary of mean AP, mean recall, mean precision

In this ROC curve for plumes, it shows the tradeoff between sensitivity and specificity, and the closer the curve follows the left-hand border and then the top border of ROC space, the more accurate the test.

The model provides regional proposal suggestions for objects within the image along with a likelihood the object belongs to a given class. Figure 5 shows a number of examples of these region proposal suggestions for classification of plumes and the other classes within the image. Most of the detections in these images are plumes, although reviewing closely you will see a detection of a light in the top right

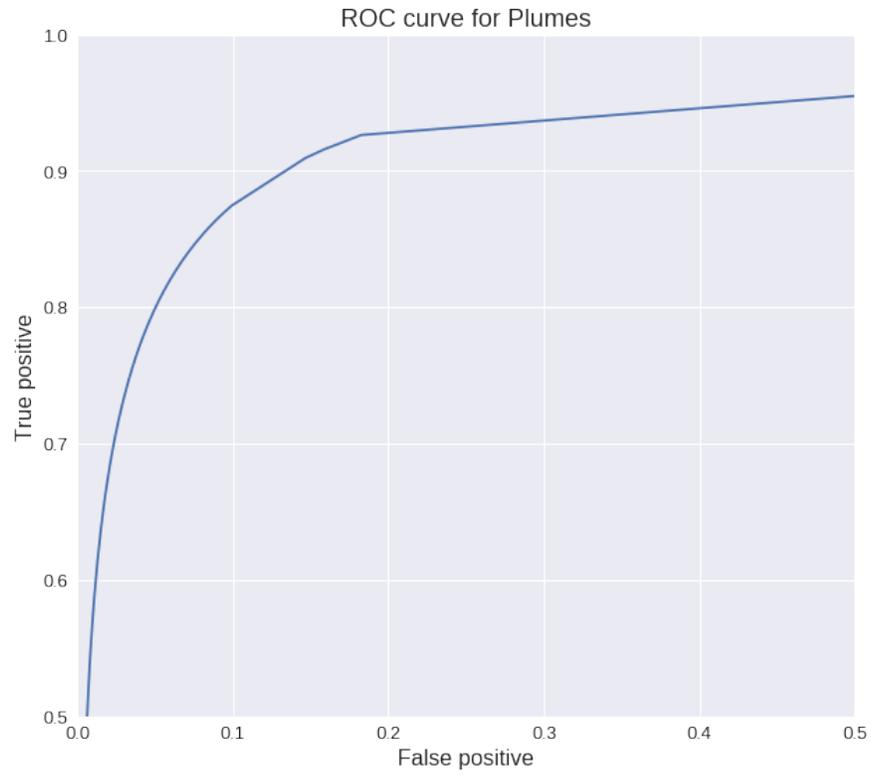


Figure 3: ROC curve for Plumes

corner of one of the images.

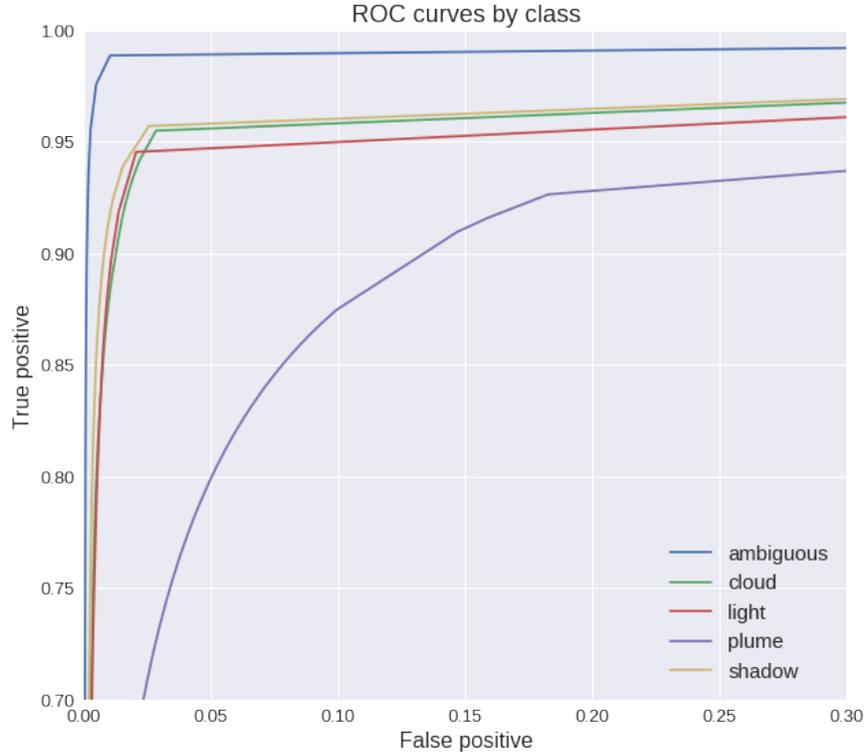


Figure 4: ROC curves by class

Figure 5: Example detection of plumes using trained model on test images. A threshold of 0.8 was applied for detections.

5 Conclusions

In this study, an image-based method for detection of polluting plumes from dense urban environments was presented. The approach fitted a regional convolutional neural network to continuously sampled images for New York city’s skyline. We showed the mean average precision of detecting plumes is 61.66 % on the testing set.

During the labeling of the training set, it was found that the temporal context was effective for the disambiguation of plumes and other plume-like patterns. It is believed that the model performance would improve by integrating that contextual information through

the use of a 3D CNN, such as Region Convolutional 3D Network (RC3D) (Xu et al. 2017), which can be used to extract spatiotemporal features capturing activities, accurately localizing the start and end times of each plume.

6 Supplementary materials

6.1 Methodology (Technical description)

6.2 Background subtraction and statistical heuristics

Plumes are difficult to identify in the original images because there are many other distracting features in the cityscape. To control for this, background subtraction methods (Figure 6), were employed in order to remove stationary objects from the frame. This works as a noise attenuation method which both subtracts out irrelevant features for detecting plumes and removes a majority of the features that are specific to this cityscape and makes the model more generalizable.

The polluting plumes emitted by buildings evolve in shape and optical depth over a couple of minutes or shorter. This characteristic evolution was exploited by showing a group of consecutive background subtracted images in order over time to identify plumes as they change in shape, expand and track the trajectory as they rise into the air. Background images were created by averaging the surrounding images and taking the difference between the target image



Figure 6: Background subtraction of the image with a plume. All stationary features are removed (shown in white) and moving objects are highlighted in black. The plume is visible slightly off to the left center of the image. Other differences in the image include the cars on the motorway, movements in the river and faint grey patches which may be light, reflection, and shadows on the buildings.

and the averaged baseline. Once subtracted from the target image, plumes are visible in the resulting frame, along with any other changes to the image such as cars, clouds, and other features which will constitute potential contaminants (false positives).

The following section provides detail on the application used for tagging of plumes and other features. We also provide a description of the different features tagged and examples in both the original and background subtracted images.

6.3 Image Labeling

In order to apply labels to the images and extract plumes to generate the training set, a web-based tool was built using d3.js, a data visualization library in Javascript [Bostock et al. \(2011\)](#), and Flask, a web application framework in Python ([Ronacher et al., 2010](#)), as is seen in Figures 7 and 8. The tool allowed us to apply bounding box

labels to batches of consecutive images; collect summary statistics on number of tags per day, or per image; linked tags between images to track appearance of the same plume across multiple images; as well as account for who applied tags to which images to assess for intrinsic human bias. The images are manipulated in Python and OpenCV (Bradski, 2000) before displaying the images on screen.

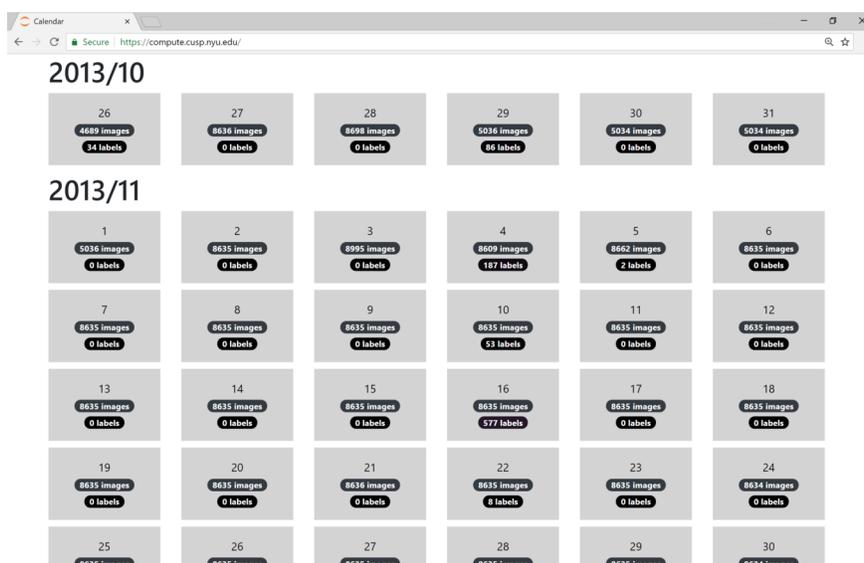


Figure 7: The plume labeling tool. Here the calendar application is displayed showing number of images per day, and number of labels which have been applied by day.

Examples of tags applied:

The tags applied included the following:

1. Plumes (two types of plumes will be outlined below)
2. Clouds
3. Shadows
4. Lights, and an

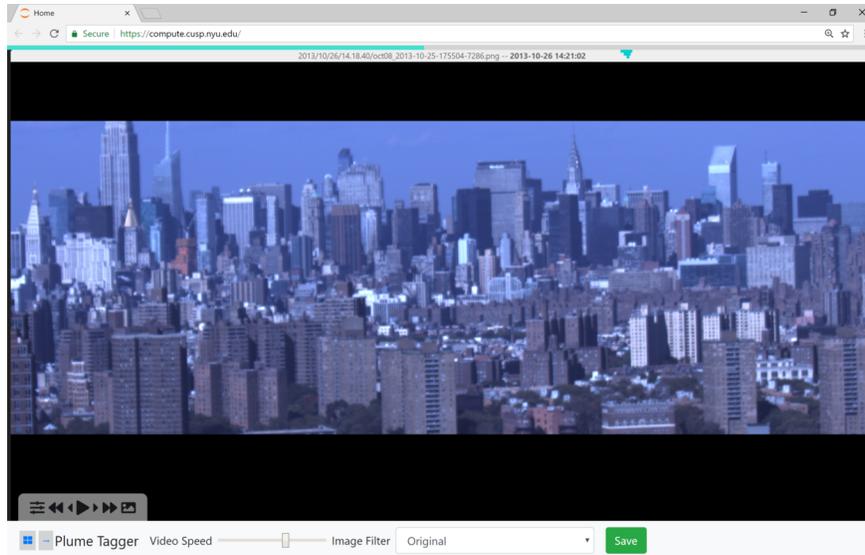


Figure 8: The user can browse the images in sequence and draw bounding boxes around instances of plumes. Here the original image is displayed. Generally, the background subtracted images will be used for labeling.

5. Ambiguous (miscellaneous) identifier

Further detail on the characteristics and features attributable to these tags is provided below.

6.3.1 Plume (black smoke cloud)

Description: A number of buildings in New York emit large black pollution plumes (as shown on the left-hand side of Figure 9). Often these are from large (and old) boilers used for building heating.

Features / Attributes: A typical example appears in around 3-10 consecutive UO images. The background subtracted images show the plume highly concentrated at the point of emission, and expanding and diffusing as the plume progresses and dissipates into the air.

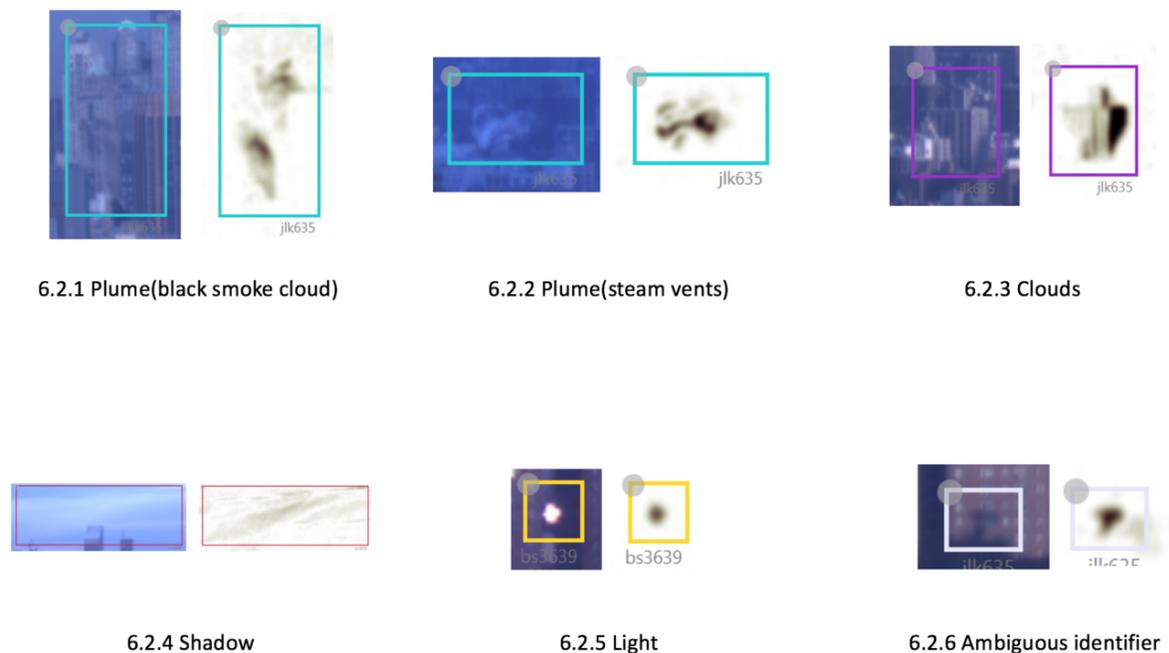


Figure 9: Example of tagged plume, cloud, shadow, light, ambiguous object shown in both the original image (left) and background subtracted image (right)

Figure 10: Animation of tagged Black smoke cloud plume in original image (The first “still image” is the frame before plume originates)

Figure 11: Animation of tagged Black smoke cloud plume in background subtracted image (The first “still image” is the frame before plume originates)

Source of contamination: Detection of black smoke plumes is highly variable across days inspected, due to weather and visibility considerations such as snow and fog. Wind is also a large determinant in visibility as high wind speeds cause faster dissipation.

Issues / challenges: Another challenge with these black smoke plumes as they rise into the air above the city skyline is they appear very similarly to clouds in the sky. The temporal context is very useful in disambiguating them, as the object trajectory is distinct from clouds.

6.3.2 Plume (steam vents)

6.3.3



Figure 12: Example of multiple steam vents tagged in single image. As shown there are 14 distinct steam vents tagged in this particular image, with the original shown on the left, and background subtracted image shown on the right.

Description: New York City has a large steam system for the purposes of heating and cooling high rise buildings and businesses. Steam vents in New York are often the result of outside water coming into contact with pipes from the steam system network; and this is more common in the winter with snow and rain.

Features / Attributes: The steam vents are seen in the UO images as a continuous small plume of grey-white cloud streaming from the tops of buildings, shown in Figure 9. As in Figure 12, there may be over 10 or more steam vents in a single image.

Source of contamination: Detection of steam vents is dependent on conditions observed over particular time periods. Variables such as wind, temperature, lighting, rain and snow can all impact both the occurrence of steam plumes, along with ability to detect these plumes in the UO images. Whilst rain and snow can make the occurrence of steam vent more common when they come into contact with the hot pipes; they also make their visibility more difficult in the images we have available.

Issues / challenges: A large number of these steam vents appear,

particularly on cold winter days. The small and continuous nature of these plumes means complete tagging is time consuming. There may also be days where there is at least one of these steam vents in every single image for the whole day.

6.3.4 Clouds

Description: Clouds are visible masses of water droplets or other particles above skyline. Level of cloud cover varies considerably, typically moving across the skyline from left to right of the image.

Features / Attributes: The slower moving clouds often do not appear in the background subtracted images because of the low rate of change in color and movement. Time periods with higher cloud activity tended to have high activity across the entire sky.

Source of contamination: For detection of polluting plumes, clouds pose a source of noise when they are dense, moving quickly and follow the same trajectory as a partially dissipated black smoke plume follows after it has risen above the city skyline. These are particular sources of signal we are interested in tagging so model can learn to distinguish as non-plumes.

Issues / challenges: The objective of tagging clouds is not to identify images containing clouds, but rather those which may serve as a source of noise for our model aimed at detecting plumes. Efforts were concentrated on tagging clouds which have similar features to the plumes of interest.

6.3.5 Shadow

Description: Shadows appear in the images as the result of light being shone, or blocked (e.g. by clouds) on part of the image over

short periods of time. High levels of shadow appear during the sunrise and sunset periods each day; and may also appear during other parts of the day dependent on weather and visibility conditions.

Features / Attributes: The shadows appear as building silhouettes in the background subtracted image. A feature we can use to distinguish shadows from plumes are the distinct edges and outlines of the buildings characteristic of the shadow when compared with the free-form characteristic of plumes.

Source of contamination: The shadows appear as a source of noise for detection of pollution plumes as they result in very dark spots and differences in the background subtracted images. This may appear like the originating source of a plume as typically the shadows are dense and can change a lot between images. The shadows can also move between images as the light or cloud moves across the image, although typically movements are horizontal across the image, rather than a rise into the air like plumes do.

Issues / challenges: Frequency and form of shadows can vary considerably across images. The objective of tagging shadows is to provide examples of inter-image motion that are not plumes to improve the model's precision.

6.3.6 Light

Description: Lights appear in the skyline at different time periods, most frequently before sunrise and after sunset. Lights are also observed on the tops of buildings and tall structures as collision avoidance measures for aircraft which may fly overhead.

Features / Attributes: The lights appear again as small, round objects in the skyline and present as bright areas in the original im-

age. The lights may appear in single images such as flashing warning lights atop building, or consecutive images when they are switched on, or off.

Source of noise: Lights are a densely concentrated object in the differenced image. They appear very similarly to when smoke clouds initiate, although they are stationary and do not dissipate in subsequent images like the pollution clouds do.

Issues / challenges: The lights comparably easy to identify for tagging, relative to clouds and shadow. They appear as a very pronounced difference in the background subtracted images, with a static shape over time.

6.3.7 Ambiguous identifier

Description: The “ambiguous” label was used to tag any miscellaneous sources of perturbation which may be confused as signal. The identifier hasn’t been used widely. Features include birds flying overhead; helicopter and aircraft; shadows of plumes on buildings (it’s not really a plume, and it’s not our typical representation of a shadow either).

Features / Attributes: The ambiguous identifier is mainly used as a catch-all for other sources of potential noise. If certain subclasses appear as false positives frequently, they may be revisited and further codified as what they actually represent (e.g. aircraft; shadow of a plume etc.)

Source of noise: These ambiguous identifiers will be used to provide extra information about miscellaneous signals that can’t be cap-

tured by the existing labels. The example shown below in Figure 9 is the shadow of a plume on a building.

Issues / challenges: The ambiguous identifiers are quite sparse and will most likely have minimal effect on model training, however they have been included where the feature appear pronounced and similar to the plumes found either elsewhere in the image or previous images.

Summary of tags applied

For compilation of the training set, images were reviewed across several days, with a goal to tag 5,000 different examples of plumes; and 1,000 different examples of clouds, lights and shadows respectively. Figure 13 below show number of plumes, clouds, lights, shadows and other categories identified (represent count of bounding boxes).

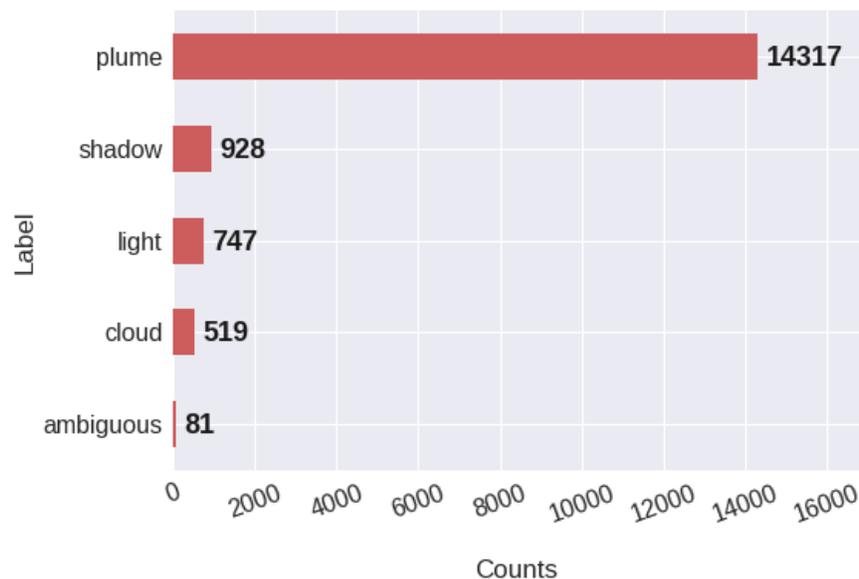


Figure 13: The number of individual bounding boxes gathered for each label. This includes separate counts for each frame that the plume appears in.

6.4 Model Architecture

The model architecture will be based on Faster R-CNN, a state-of-the-art object detection CNN that uses a region proposal algorithm to hypothesize object locations (Ren et al., 2017). CNN uses a series of pattern detectors that the model learns from training data, and classify images in ImageNet training set into the different classes (Krizhevsky et al., 2017). In order to derive object location information, R-CNN (the antecedent to Fast R-CNN and Faster R-CNN) creates region proposals, using a process called Selective Search (Girshick et al., 2014), and run the images in bounding boxes through a pre-trained AlexNet and finally use SVM to classify objects. To speed up and simplify R-CNN, Ross Girshick came up Fast R-CNN with RoI (Region of Interest) Pooling, which shares the forward pass of a CNN for an image across its subregions (Girshick, 2015). Based on the works of CNN, Faster R-CNN was chosen, as it runs more quickly and efficiently by reusing the convolutional features for both the image classification and the region proposal (Ren et al., 2017).

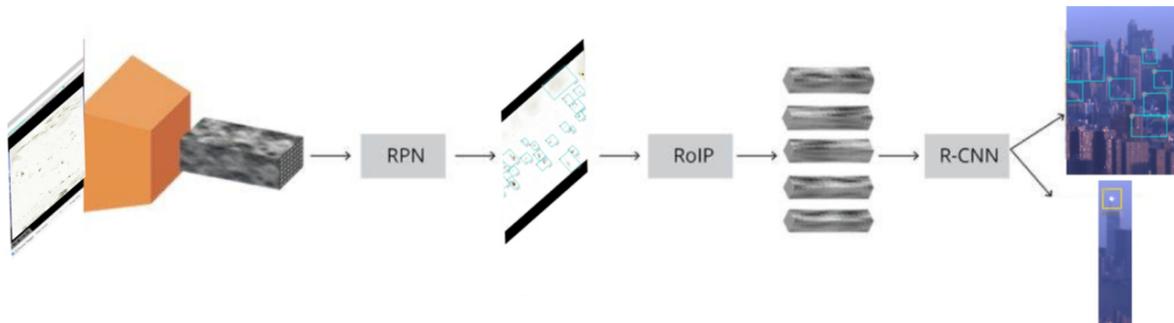


Figure 14: Complete Faster-RCNN architecture used for detection of pollution plumes in UO images.

The detailed model architecture was developed and tested on the PASCAL Visual Object Classes (VOC) dataset. [Ren et al. \(2017\)](#) included existing image sets, development kit functions, classification functions, detection functions, segmentation functions and layout functions. The Faster R-CNN model will be used as the foundation for training the model for detecting the plumes, and weights pre-trained on the VOC 2007 dataset will be used to initialize the model. The dataset used was converted into the PASCAL VOC data annotation format for easier compatibility with the existing code base.

6.5 Plume detection census

An important piece of information to extract from instances of plumes is the building that the plume originated from. This is a non-trivial task because it requires reconstructing 3D information from using only a 2D image. Photogrammetry is a field of research that is used to undertake this, and an existing mapping from image pixel to the 3-dimensional location was used to geolocate the plume with respect to its building of origin.

7 Discussion

Visual identification of instances of plumes is performed over a subset of the dataset in order to generate a labeled dataset for training the model. Due to the large number of images, the process of labeling was very time-consuming. In order to improve the efficiency of the process, statistical heuristics were used to find points of separation between images that contained plumes and those that do not. This

approach whilst useful, was not always conclusive and one aspect to be wary of were the selection biases associated with any of these techniques. Other challenges associated with labeling the dataset included variations in weather and visibility patterns across different time periods which made plumes difficult to identify. Additionally to this, usage patterns vary throughout the time periods meaning that frequency, volume, and type of plumes varied considerably. We needed to be conscious of this when identifying the training set to ensure a suitably representative set of images were reviewed (by multiple different reviewers) to tag enough plumes across different conditions such that the model will be more robust to environmental variations.

References

- M. Bostock, V. Ogievetsky, and J. Heer. *D³Data—Driven Documents*. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, dec2011. doi: . URL <https://doi.org/10.1109/2Ftvcg.2011.185>.
- Gary Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- Juliane Brink and Edzer Pebesma. Plume Tracking with a Mobile Sensor Based on Incomplete and Imprecise Information. *Transactions in GIS*, 18(5):740–766, sep 2013. 10.1111/tgis.12063. URL <https://doi.org/10.1111/2Ftgis.12063>.
- City of New York. PLANYC 2030 - Greener, Greater Buildings Plan. 2010. URL <http://www.nyc.gov>.
- City of New York. New York City Local Law 84 Benchmarking Report., 2013. URL <http://www.nyc.gov/html/gbee/downloads/>

[pdf/nyc_energy_water_use_2013_report_final.pdf](https://www.sciencedirect.com/science/article/pii/S0306437915001167).

Gregory Dobler, Masoud Ghandehari, Steven E.Koonin, Rouzbeh Nazari, Aristides Patrinos, Mohit S.Sharma, Arya Tafvizi, Huy T.Vo, and Jonathan S.Wurtelec. Dynamics of the urban lightscape. *Information Systems*, Volume 54 Issue C, 2015. URL <https://www.sciencedirect.com/science/article/pii/S0306437915001167>. Accessed on Sun, April 29, 2018.

Sebastien Frizzi, Rabeb Kaabi, Moez Bouchouicha, Jean-Marc Ginoux, Eric Moreau, and Farhat Fnaiech. Convolutional neural network for video fire and smoke detection. In *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE, oct 2016. 10.1109/iecon.2016.7793196. URL <https://doi.org/10.1109%2Fiecon.2016.7793196>.

Ross Girshick. Fast R-CNN. In *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, dec 2015. 10.1109/iccv.2015.169. URL <https://doi.org/10.1109%2Ficcv.2015.169>.

Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2014. 10.1109/cvpr.2014.81. URL <https://doi.org/10.1109%2Fcvpr.2014.81>.

Jayavardhana Gubbi, Slaven Marusic, and Marimuthu Palaniswami. Smoke detection in video using wavelets and support vector machines. *Fire Safety Journal*, 44(8):1110–1115, nov 2009. 10.1016/j.firesaf.2009. URL <https://doi.org/10.1016%2Fj.firesaf.2009.08.003>.

Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. Large-Scale Video Classifica-

tion with Convolutional Neural Networks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2014. 10.1109/cvpr.2014.223. URL <https://doi.org/10.1109%2Fcvpr.2014.223>.

ByoungChul Ko, JunOh Park, and Jae-Yeal Nam. Spatiotemporal bag-of-features for early wildfire smoke detection. *Image and Vision Computing*, 31(10):786–795, oct 2013. 10.1016/j.imavis.2013.08.001. URL <https://doi.org/10.1016%2Fj.imavis.2013.08.001>.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, may 2017. 10.1145/3065386. URL <https://doi.org/10.1145%2F3065386>.

Nicholas L Mills, Ken Donaldson, Paddy W Hadoke, Nicholas A Boon, William MacNee, Flemming R Cassee, Thomas Sandström, Anders Blomberg, and David E Newby. Adverse cardiovascular effects of air pollution. *Nature Clinical Practice Cardiovascular Medicine*, 6(1):36–44, nov 2008. 10.1038/ncpcardio1399. URL <https://doi.org/10.1038%2Fncpcardio1399>.

NYC Health. New York City Trends in Air Pollution and its Health Consequences, 2013. URL <https://www1.nyc.gov/assets/doh/downloads/pdf/environmental/air-quality-report-2013.pdf>.

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, jun 2017. 10.1109/tpami.2016.2577031. URL <https://doi.org/10.1109%2Ftpami.2016.2577031>.

Armin Ronacher, David Lord, Adrian Mönnich, and Marcus Unter-

waditzer. Flask (A Python Microframework). *Flask*, 1, 2010.

Huijuan Xu, Abir Das, and Kate Saenko. R-C3D: Region Convolutional 3D Network for Temporal Activity Detection. *Proceedings of the International Conference on Computer Vision (ICCV)*, 2017. URL <https://arxiv.org/abs/1703.07814>.

Feiniu Yuan. A fast accumulative motion orientation model based on integral image for video smoke detection. *Pattern Recognition Letters*, 29(7):925–932, may 2008. 10.1016/j.patrec.2008.01.013. URL <https://doi.org/10.1016%2Fj.patrec.2008.01.013>.

Turgay Çelik, Hüseyin Özkaramanlı, and Hasan Demirel. Fire And Smoke Detection Without Sensors: Image Processing Based Approach. In *Computer Engineering and Intelligent Systems*. EURASIP, 2007. URL <https://www.eurasip.org/Proceedings/Eusipco/Eusipco2007/Papers/c5p-j09.pdf>.